

# Using GORE in Big Data: A Systematic Mapping Study

A. Cravero and S. Sepúlveda

**Abstract**—Big Data has developed rapidly into a hot topic that generates great attention in academia, industry and governments around the world because of its ability to process large volumes of data at a reasonable speed. The real value of Big Data lies in the help to business decision making, so it requires appropriate methods and techniques to find the requirements of the system to be developed. However, the Big Data community has focused mainly on technologies to obtain information, and not on the goal of the business. The objective of this paper is to offer an overview of how some Goal Oriented Requirements Engineering (GORE) proposals are used to obtain the requirements of the Big Data system, serving as a starting point for future research. The methodology used is the Systematic Mapping. The goal models identified it is used as conceptual models for different situations: understanding the nature of business and analyzing the technical aspects for purposes of functional and non-functional requirements of Big Data system.

**Index Terms**—Big Data, GORE, Goal-oriented, Requirements, Systematic mapping.

## I. INTRODUCCIÓN

ES evidente la influencia de la orientación a objetivos (conocido como Goal-Oriented Requirement Engineering o GORE) en los métodos y técnicas de Ingeniería de Requisitos (IR) contemporáneos, debido a que un gran número de métodos de IR lo utilizan como un medio de abstracción de alto nivel para estructurar y abstraer el contenido de los requisitos del sistema [1].

La Ingeniería de Requisitos Orientada a Objetivos (GORE por sus siglas en inglés) aporta varios beneficios a la práctica de IR, tales como: criterios precisos para la completitud suficiente de una especificación de requisitos; justifica la existencia de un requisito; permite estructurar documentos formales de requisitos complejos aumentando la comprensión; permite detectar y resolver conflictos entre los requisitos; entre otros [2]. De esta manera se puede distinguir las necesidades de requisitos en las etapas tempranas de IR versus las tardías (educación, especificación, negociación y validación) debido a que puede conducir a concepciones de

objetivos y métodos de análisis de objetivos que son sustancialmente diferentes [3].

Varios autores concluyen que es crucial el uso de GORE debido a la falta de comprensión de los objetivos estratégicos por parte de los analistas de sistemas [2, 4-8]. El beneficio principal es el alineamiento entre las Tecnologías de Información (TI) y los objetivos estratégicos de la organización.

A pesar de los beneficios de GORE, existe poca evidencia de su uso en los sistemas de Big Data. Big Data se ha desarrollado rápidamente en un tema candente que genera gran atención en la academia, la industria y los gobiernos de todo el mundo debido a su posibilidad de procesar grandes volúmenes de datos a una velocidad razonable [9].

Hasta ahora, la comunidad de Big Data se ha centrado en los requisitos tecnológicos de su implementación, y no en los que aportan a la comprensión de los objetivos del negocio [10]. En este sentido, el hardware para Big Data tiene requisitos más estrictos sobre la capacidad de almacenamiento y la capacidad de procesamiento, así como la capacidad de transmisión de la red [11].

Por otro lado, las herramientas analíticas actuales, como Big Data, no prestan suficiente atención a la evaluación de los objetivos del negocio o la mitigación de los problemas comerciales a través de los conocimientos derivados de los datos. Este problema se reconoce en una encuesta industrial global de casi 3000 ejecutivos, gerentes y analistas de todo el mundo [10].

Otra encuesta realizada en Julio de 2012 a 500 empresas del área financiera, con la participación de más de 50 ejecutivos senior que completaron un conjunto muy detallado de 65 preguntas, entregó como principal resultado que la clave para la implementación de proyectos de Big Data es la alineación organizacional, catalogándolo como un factor muy crítico para garantizar el éxito [12].

La construcción de un sistema de Big Data no es una tarea fácil, en comparación con la ingeniería de software, es una disciplina muy joven y aún no existen estrategias y técnicas bien establecidas para el proceso de desarrollo [13]. En este sentido GORE permite mejorar el diseño de la solución de Big Data para la organización, debido a que los modelos de objetivos permiten el alineamiento entre la TI y el plan estratégico empresarial [14, 15].

Hasta el momento, no existe una estrategia común para el desarrollo de sistemas de Big Data. Pero en los últimos años, se han propuesto algunos métodos y técnicas que permiten a

Este trabajo ha sido financiado por la Universidad de La Frontera a través del proyecto DIUFRO DI17-0099.

A. Cravero, es Directora del Departamento de Cs. de la Computación e Informática, Universidad de La Frontera, Temuco, Chile.  
(email: ania.cravero@ufroterra.cl).

S. Sepúlveda, es Académico del Departamento de Cs. de la Computación e Informática, Universidad de La Frontera, Temuco, Chile.  
(email: samuel.sepulveda@ufroterra.cl).

los desarrolladores obtener los requisitos de información desde el modelo de objetivos empresarial [16]. Es por ello que el objetivo de este artículo es ofrecer una visión general de cómo algunas propuestas utilizan GORE para obtener los requisitos del sistema de Big Data, sirviendo como punto de partida para futuras investigaciones.

La metodología utilizada para la selección de los trabajos es el mapeo sistemático. El mapeo sistemático de estudios es una metodología utilizada frecuentemente en la investigación médica, la que ha sido adaptado para su uso en el área de TI [17].

El documento se estructura de la siguiente manera. La sección II presenta los conceptos básicos utilizados en el trabajo. La sección III describe el proceso de selección de los estudios. La sección IV describe el análisis de los estudios con respecto al uso de GORE en los sistemas de Big Data. Finalmente, la sección V concluye el estudio.

## II. CONCEPTOS BÁSICOS

### A. Big Data

Existen distintas definiciones aportadas por los estudios del tema de Big Data. Entre las cuales se tiene, una nueva generación de tecnologías y arquitecturas diseñadas para extraer económicamente el valor de volúmenes muy grandes de una amplia variedad de datos al permitir la captura de alta velocidad, el descubrimiento y/o análisis [18]. Una definición incluye las propiedades conocidas como las 5V: Volumen, Velocidad, Variedad, Valor y Veracidad [19]. Explican que Big Data no es sólo una base de datos o un problema Hadoop, sino que es todo el conjunto de componentes, o arquitectura, para almacenar, procesar, visualizar y entregar resultados para orientar las aplicaciones. Según los mismos autores, este conjunto de componentes interrelacionados se puede definir como el Ecosistema Big Data que se ocupa de la evolución de los datos, los modelos e infraestructura de apoyo durante todo el ciclo de vida de Big Data [20]. Otras definiciones se pueden encontrar en [21, 22].

Como resultado del análisis de las prácticas existentes en las diferentes comunidades científicas y dominios tecnológicos de la industria, se ha propuesto el modelo del Big Data Lifecycle Management (BDLM) (ver Fig. 1), que es el enfoque necesario para la gestión y procesamiento de los datos masivos [19].

La Fig. 1 describe el proceso general que se debe llevar a cabo a la hora de crear un proyecto de Big Data. El proceso se inicia con la identificación de las fuentes desde donde se extraerán los datos útiles. A continuación, los datos son almacenados en alguno de los modelos de datos diseñados, de acuerdo a si son datos estructurados o no. En el siguiente paso, los datos se clasifican y se filtran de acuerdo al tipo de análisis que se desee realizar. Los datos clasificados son analizados utilizando herramientas adecuadas para ello; ejemplo de ello, son las de minería de datos, las de OLAP, y las de ciencias en general. Los datos obtenidos deben ser presentados con alguna herramienta de visualización. Finalmente los datos son analizados por los responsables de tomar decisiones [20].

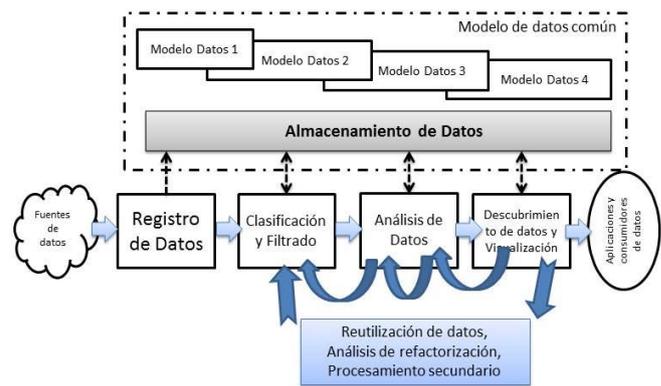


Fig. 1: Ciclo de vida del Ecosistema de Big Data [19].

### B. Ingeniería de Requisitos Orientada a Objetivos – GORE

La calidad de un sistema de software depende fundamentalmente del grado en que cumpla con sus requisitos. Dichos requisitos se pueden obtener, modelar y analizar como objetivos de los interesados [4]. GORE surgió para crear y estudiar diversos métodos que abordan la IR desde una perspectiva orientada a objetivos. Normalmente, dentro de GORE, los objetivos se obtienen y conceptualizan en términos de alguna forma de modelo [23]. Los modelos de objetivos se han utilizado como un medio eficaz para capturar las interacciones y las compensaciones entre los requisitos, pero se han aplicado de manera más amplia para avanzar en el estado de la adaptación del software, la seguridad, el cumplimiento legal y la inteligencia comercial, entre otras áreas [4]. Estudios recientes muestran la evolución de GORE, y la forma cómo se está utilizando en los distintos tipos de sistemas [24, 25].

Hoy en día los sistemas de apoyo a la decisión responden a una parte importante de los presupuestos TI en la mayoría de las organizaciones [26]. A pesar de los beneficios, muchos proyectos no pueden entregar la información esperada para apoyar la toma de decisiones del negocio [27, 28], debido a que falta rigurosidad por parte de los analistas de sistemas para utilizar estrategias de GORE que permitan obtener requisitos más pertinentes al contexto del negocio [29]. A pesar de disponer de toda la experiencia en el desarrollo de Almacenes de Datos, aún se sigue teniendo los mismos problemas [29]:

- 1) Falta de alineamiento entre TI y el plan estratégico de la organización, que es la clave para el éxito del proyecto [12].
- 2) Se requiere de personas especialistas con las habilidades adecuadas [12].

En este sentido, el objetivo primordial de la etapa de IR para un sistema de Big Data, es representar las necesidades de los usuarios, y los objetivos del sistema, con el fin de alcanzar los objetivos estratégicos del negocio [30]. Esta etapa es crucial, debido a que permite a los analistas de sistemas crear aplicaciones de Big Data que tengan sentido para la organización alineándolo con los objetivos estratégicos planteados.

Esta es la razón por que la etapa de IR para sistemas de Big Data debe basarse en el marco de GORE propuesto por Kavakli, debido a que (1) Big Data tiene por objetivo proporcionar información adecuada para apoyar la toma de decisiones, contribuyendo así a cumplir los objetivos estratégicos de una organización, (2) los requisitos son difíciles de obtener a partir de cero, debido a que los ejecutivos a menudo sólo expresan las expectativas generales [3]. En este sentido la introducción de objetivos ofrece una forma de clarificar los requisitos de un sistema [31].

El marco identifica cuatro tareas que deben ser utilizadas en distintos contextos. (1) La educación de requisitos, que se enfoca en la descripción de las necesidades de la organización y las limitaciones del sistema a desarrollar; (2) la negociación, que permite establecer un acuerdo sobre los requisitos del sistema entre los diferentes actores involucrados en el proceso; (3) la especificación de requisitos, que pretende entender la situación actual de la organización y las consideraciones para mejorar el sistema, y por último; (4) la tarea de validación de requisitos, que debe garantizar que los requisitos derivados corresponden a las necesidades originales de la organización y al ajuste con las necesidades externas fijadas por la empresa y su entorno. De acuerdo a estas tareas identificadas, Kavakli clasifica los diferentes modelos de objetivos que son utilizados en la GORE [3] (ver Tabla I).

TABLA I  
CONTRIBUCIÓN DE LOS MODELOS DE OBJETIVOS SEGÚN LA ETAPA DE IR BASADO EN [33]

Actividades de IR	Contribución de las aproximaciones	Aproximaciones Orientadas a Objetivos
<i>Edución</i>	1. Comprensión de la situación actual de la organización, 2. Comprensión de la necesidad de cambio.	GOMS, Goal-based Workflow, i*-Tropos, ISAC, F3, EKD, AGORA
<i>Negociación</i>	3. Proporcionar el contexto del proceso de IR.	SIBYL, The reasoning loop model
<i>Especificación</i>	4. Relacionar los objetivos del negocio con los componentes del sistema.	KAOS, GBRAM, NFR, GSC
<i>Validación</i>	5. Validación de las especificaciones del sistema en contra de los objetivos de las partes interesadas.	GQM, GSM, URN, AGORA

Kavakli argumenta que existe una distinción entre las necesidades previstas en etapas tempranas en relación con las etapas posteriores que pueden llevar a distinguir distintos aspectos de los objetivos [32]. Durante las primeras etapas de la GORE es más importante modelar y analizar las necesidades de los actores (en nuestro caso, ejecutivos o responsables de tomar decisiones) y sus intereses y cómo podrían ser comprometidos por la decisión de introducir un nuevo sistema [3]. Las etapas posteriores se refieren a los objetivos futuros y cómo estos pueden llevarse a la práctica en términos de componentes del sistema.

C. Mapeo Sistemático de Estudios

La técnica de mapeo sistemático (o systematic mapping) define un proceso y una estructura de informe que permite categorizar los resultados que han sido publicados hasta el momento en un área determinada [34].

El objetivo de un mapeo sistemático está en la clasificación, y está por tanto dirigido al análisis temático y a la identificación de los principales foros de publicación. Permite responder preguntas genéricas como ¿Qué es lo que se ha hecho hasta el momento en el campo X? Como limitación, este tipo de estudios no toma en consideración la calidad de los estudios incluidos.

El proceso de mapeo sistemático consiste en las siguientes etapas: (1) definición de las preguntas de investigación, (2) ámbito de la revisión, (3) ejecución de la búsqueda, (4) selección de estudios, (5) filtrado de estudios, (6) esquema de clasificación, (7) extracción de datos, (8) y mapa sistemático [34].

III. PLANTEAMIENTO DEL PROBLEMA Y ACTIVIDADES DESARROLLADAS EN EL MAPEO SISTEMÁTICO

A continuación, se describe las actividades principales desarrolladas en el mapeo sistemático.

A. Definición de las Preguntas de Investigación

Las siguientes Preguntas de Investigación (PI) se definieron de acuerdo a la técnica indicada por [35]:

PI1: ¿Qué modelos de objetivos son utilizados en los sistemas de Big Data y en qué sub-etapa de IR?

PI2: ¿Los modelos de objetivos utilizados por los estudios, coinciden con la clasificación en el marco de GORE propuesto por Kavakli?

PI3: ¿Cuál es la contribución de los trabajos de investigación seleccionados? (si es metodología, enfoque, técnica, herramienta)

PI4: ¿Cuál es el ámbito de aplicación del estudio? (si es en la academia, o industria)

PI5: ¿Para qué son utilizados los modelos de objetivos en sistemas de Big Data?

PI6: ¿Qué resultados reportan los estudios seleccionados?

Para quienes se interesan en la investigación, se proponen dos hipótesis que deben ser respondidas dentro de las preguntas anteriores.

- 1) Hipótesis 1 “GORE mejora el proceso de IR para sistemas de Big Data debido a que introduce el punto de vista del negocio en los requisitos de análisis de datos”
- 2) Hipótesis 2: “es posible adaptar los modelos de objetivos utilizados en otros contextos, en los sistemas de Big Data”

Estas hipótesis permitirán a los futuros investigadores decidir si es conveniente continuar con nuevas investigaciones de GORE en los sistemas de Big Data, además de identificar las opciones de su uso.

### B. Ejecución de la Búsqueda

La cadena de búsqueda consistió en expresiones booleanas formados por palabras claves: ("big data" OR "big-data") AND ("goal-oriented" OR "GORE" OR "goal"). Para la búsqueda fue importante que la palabra clave "big data" estuviera contenido en el título de cada estudio, por otro lado GORE o goal-oriented o goal deben ser conceptos encontrados en las palabras claves definidas por el autor.

Las fuentes donde se aplicó la búsqueda son: Scholar Google, Scopus, IEEE Digital Library, ACM Digital Library y Springer.

### C. Selección de Estudios y Proceso de Filtrado

La selección de los estudios se ha formulado basada en los siguientes criterios de inclusión/exclusión.

#### 1) Inclusión

Trabajos de investigación que provienen de revistas y congresos, y que describan técnicas, aproximaciones, metodologías y/o herramientas para obtener los requisitos del sistema de Big Data a partir de un modelo de objetivos.

#### 2) Exclusión

Documentos, y trabajos que tratan sobre el uso de modelos de objetivos, en cualquiera de sus temas de interés; y trabajos que se centran en el desarrollo de sistemas de Big Data, pero que no explican el uso de modelos de objetivos para obtener requisitos.

El proceso de selección consta de dos iteraciones realizadas por dos colaboradores. En la primera iteración, cada revisor aplicó los criterios de inclusión y exclusión para el título, resumen y palabras clave para los 18 estudios obtenidos en Scholar Google (ver fila 1 en la tabla II), de manera separada. Una vez obtenida la lista de estudios, los revisores compararon los resultados con el fin de asegurar que ambos aplicaron dichos criterios con rigurosidad (ver resultados en fila 2). Para el caso de los trabajos que generaba dudas, los revisores debieron leer la introducción y conclusiones para luego debatir su pertinencia para el trabajo. De esta manera se obtuvo un total de 5 estudios (fila 3), 4 de ellos se pueden encontrar en Scholar Google, 1 de ellos en Scopus y 3 de ellos también se encuentran en IEEE. No se encontraron estudios en ACM y Springer.

TABLA II  
RESULTADOS DE LA BÚSQUEDA Y FILTRADO

Motor de Búsqueda	IEEE	ACM	Springer	Scopus	Scholar Google
Resultados de la búsqueda	5	8	6	16	18
Candidatos	3	2	2	5	6
Trabajos seleccionados	3	0	0	1	5

La Tabla III, presenta la lista de los estudios seleccionados de acuerdo a los criterios de inclusión y exclusión.

Por otro lado, la Tabla IV presenta la lista de estudios relacionados con Big Data y el uso de objetivos del negocio para obtener requerimientos. Cabe destacar que estos estudios no utilizan un modelo de objetivos, por ello no son incluidos para dar respuesta a las preguntas de investigación.

TABLA III  
ARTÍCULOS SELECCIONADOS

Autores	Título	Año	Tipo publicación	Revista o Conferencia
H. Eridaputra, B. Hendradjaya	Modeling the requirements for big data application using goal oriented approach	2014	Conferencia	ICODSE
S. Supakkul, L. Zhao, L. Chung	GOMA: Supporting Big Data Analytics with a Goal-Oriented Approach	2016	Conferencia	Big Data Congress
D. Fekete	The GOBIA Method: Fusing Data Warehouses and Big Data in a Goal-Oriented BI Architecture	2016	Conferencia	GvD
G. Park, S. Park, L. Khan, L. Chung	IRIS: A goal-oriented Big Data analytics framework on Spark for better Business decisions	2017	Conferencia	BigComp
M. Fahmideh, G. Beydoun	Big data analytics architecture design—An application in manufacturing systems	2018	Revista	Computers & Industrial Engineering

TABLA IV  
ARTÍCULOS RELACIONADOS

Autores	Título	Año	Tipo publicación	Revista o Conferencia
Rapp, Geoffrey Christopher	Intelligence Design: An Analysis of the SEC's New Office of Market Intelligence and Its Goal of Using Big Data to Improve Securities Enforcement	2013	Revista	U. Cin. L. Rev.
Mao-hua, Ma and Qian, Chen	Goal Orientation and Path Selection of Libraries' Digital Humanities Construction in Big Data Era	2017	Revista	Hebei Library Journal of Science and Technology
Boyd, Stephen and Parikh, Eric and Neal and Chu, Eric and Peleato, Borja and Eckstein, Jonathan and others	Distributed Framework for Big Data Processing: a Goal Driven Approach	2011	Revista	Foundations and Trends in Machine learning
Golfarelli, Matteo and Pirini, Tommaso and Rizzi, Stefano	Goal-Based Selection of Visual Representations for Big Data Analytics	2017	Conf.	International Conference on Conceptual Modeling. Springer
Gray, Chris Hables	Big Data, Actionable Information, Scientific Knowledge and the Goal of Control	2014	Revista	Teknokultura
Gabay, Clive and Ilcan, Suzan	The Affective Politics of the Sustainable Development Goals: Partnership, Capacity-Building, and Big Data	2017	Revista	Globalizations. Taylor & Francis

Kharrazi, Ali and Qin, Hua and Zhang, Yi	Urban big data and sustainable development goals: Challenges and opportunities	2016	Revista	Sustainability. Multidisciplinary Digital Publishing Institute
Sliman, Layth and Charroux, Benoit and Stroppa, Yvan	Distributed framework for big data processing: A goal driven approach	2014	Revista	Smart Digital Futures. IOS Press

Cabe destacar que hay trabajos muy interesantes que describen el uso de métodos, técnicas y modelos (no de objetivos) de IR que ayudan a enfrentar el diseño conceptual de arquitecturas de Big Data desde el punto de vista de los usuarios, y desde el punto de vista del negocio. En este sentido, Lau aplica un método de IR desde la perspectiva social, basándose en la comprensión a priori, a través de teorías y modelos, de cómo los humanos dan sentido a los datos. El modelo de sensaciones que obtiene permite el diseño de una arquitectura conceptual para comprender los requisitos del usuario y las características del sistema del análisis de Big Data [36].

Por otro lado, Arruda lleva a cabo una búsqueda sistemática en la que explica que actualmente, no se conocen modelos de artefactos de IR que apoyen el diseño del proceso de IR y la comprensión de proyectos de Big Data. Para llenar este vacío, propone un método que incluye un modelo de IR para aplicaciones de usuario final de Big Data [37]. En un estudio posterior [38], explica que cuando se trata de abordar los objetivos del negocio en el proceso de desarrollo de software de Big Data, se han realizado pocas investigaciones sobre este tema. Algunos de los enfoques que describe son: (i) un framework de análisis para respaldar la toma de decisiones y la alineación con el negocio [39] (ii) un enfoque para capturar objetivos de negocio que serán usados en el análisis [10], (iii) el uso de metodologías ágiles junto con análisis de Big Data para abordar los objetivos del negocio [40], (iv) enfoque para monitorear los datos operativos que determina si una empresa está bien encaminada en relación con sus objetivos estratégicos [41], (v) enfoque para integrar sistemáticamente información descriptiva y predictiva (resultados de Big Data analítica) en la toma de decisiones para ayudar a alcanzar los objetivos comerciales [42]. Estos estudios analizan los aspectos analíticos de Big Data e intentan alcanzar los objetivos del negocio definidos mediante el uso adecuado de Big Data Analytics para tomar decisiones informadas. Arruda concluye que los enfoques emergentes actuales que se ocupan de la IR para sistemas Big Data tienden a tratar los objetivos del negocio como un problema secundario [38].

Otro autor que aborda el tema de alineamiento es Madhavji, [13] quien presenta un modelo de contexto de Big Data Software Engineering (BDSE) que contiene diferentes elementos, como la práctica de desarrollo, los sistemas de Big Data, la toma de decisiones corporativas y la investigación, y sus relaciones. Los autores también introdujeron algunos de los desafíos de investigación en ingeniería de requisitos en el

contexto de las aplicaciones de Big Data.

Chen [43] propuso un método de diseño de Big Data System, llamado BDD (Big Data-system Design), que incorpora los objetivos de negocio, las restricciones, los escenarios de calidad y los escenarios de arquitectura de Big Data.

Por otro lado Nalchigar, propone un marco de modelado (que incluye un conjunto de metamodelos y un conjunto de catálogos de diseño) para el análisis de requisitos y el diseño de sistemas de análisis de datos [44]. A pesar que el marco no es utilizado para el diseño de Big Data, sí puede ser utilizado desde la perspectiva del análisis. Así también Ardagna crea un proceso para obtener y clasificar los requisitos de una plataforma de Big Data como servicio [45]. Su trabajo proporciona una visión general del análisis realizado sobre los requisitos generales relacionados con los objetivos del proyecto, y los aspectos legales de una organización.

#### D. Esquema de Clasificación

El esquema de clasificación se crea a partir de las preguntas de investigación y los conceptos claves que se obtienen de las mismas. Los conceptos han sido clasificados como sigue:

- 1) *Ámbito de aplicación:* El área donde se desarrolla la investigación, o donde apuntan los autores para aplicar su investigación, que puede ser la academia o la industria. Academia, esta clasificación es para aquellas publicaciones que dirigen su esfuerzo en realizar nuevas investigaciones y/o desarrollo de nuevas ideas. Industria, en este estudio esta clasificación corresponde a los trabajos que aplican su investigación en alguna organización (con o sin fines de lucro).
- 2) *Modelos de objetivos:* Se refiere al modelo de objetivos utilizado para el desarrollo del sistema de Big Data. Puede ser cualquiera de los clasificados por Kavakli en la Tabla I. Si el modelo utilizado no se encuentra en la Tabla I, será clasificado como Otro.
- 3) *Etapas de IR:* Puede ser cualquiera de las etapas de IR que explica Kavakli. Educación, Negociación, Especificación y Validación.
- 4) *Tipo de contribución:* El aporte que realiza la investigación al área, que puede ser una metodología, técnica o herramienta. Metodología, incluye descripciones y procedimientos a seguir para obtener los requerimientos desde un modelo de objetivos. Técnica, se refiere a recursos o habilidades utilizadas para obtener el modelado de objetivos y los requisitos del sistema. Herramienta, se refiere a cualquier tipo de herramienta o software que ayude en el proceso de modelado y obtención de requisitos.
- 5) *Tipo de Uso:* Para este estudio, Implementación, se refiere al uso de GORE como análisis y diseño para sistemas de Big Data; Integración, se refiere al uso GORE en sistemas que integran a otros sistemas o plataformas; y Extensión, cuando el modelo de objetivos es usado para agregar nuevos componentes al modelo.

La Fig. 2 presenta el esquema de clasificación completo.

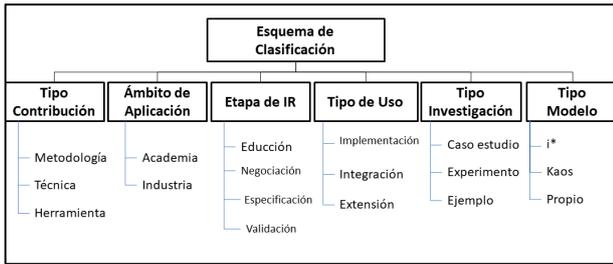


Fig. 2. Esquema de Clasificación.

E. Extracción de datos y Mapeo Sistemático

Tras definir el esquema de clasificación, el último paso del mapeo sistemático consiste en la extracción de datos y el proceso de mapeo de las distintas dimensiones. El resultado completo de esta actividad se muestra en la siguiente sección. Así, la síntesis del estudio se puede observar de manera gráfica en el diagrama de burbuja de la Fig. 3, que ilustran un diagrama de dispersión XY con burbujas en las intersecciones de categoría, permitiendo tener en cuenta varias categorías al mismo tiempo y entregar una visión general y rápida de un campo de estudio, proporcionando un mapa visual.

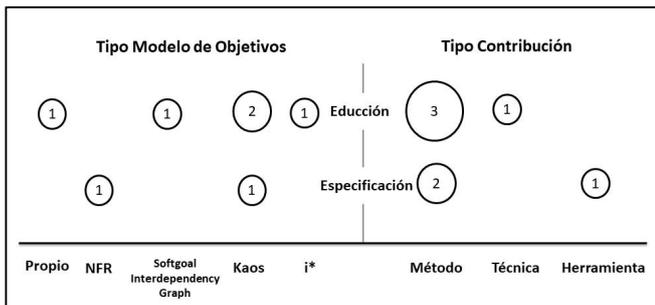


Fig. 3. Mapa Sistemático. Visualización según el tipo de modelo de objetivos y etapa de IR versus tipo de contribución.

En esta visualización de los resultados, el tamaño de una burbuja es proporcional al número de artículos que están en el par de categorías que correspondan a las coordenadas en las que se ubica la burbuja. Se puede observar la distribución de los trabajos por tipo de modelo de objetivos utilizado, y por tipo de contribución, en la etapa de IR correspondiente. Cabe destacar que hay estudios que utilizan el modelo de objetivos para educación y también para la especificación.

IV. RESULTADOS, ANÁLISIS COMPARATIVO Y DISCUSIÓN

A continuación damos respuesta a las preguntas de investigación formuladas en la sección III a través de los resultados obtenidos.

A. P1. ¿Qué modelos de objetivos son utilizados en los sistemas de Big Data y en qué sub-etapa de la IR?

Desde la Fig. 3 observamos que son pocos los modelos de objetivos utilizados para obtener requisitos para los sistemas de Big Data. La mayoría de éstos son usados para la educación de requisitos, y sólo dos para la especificación.

Tanto i\* como Kaos, han sido utilizados para la educación de requisitos. Llama la atención que dos de los trabajos utilicen

modelos de objetivos propios o adaptados. Por un lado se utilizan rectángulos para representar los objetivos del negocio que serán usados para obtener los requisitos del sistema [29]. Por otro lado, utilizan los Softgoal de i\* para representar los objetivos del negocio. Este tipo de modelo lo llaman Softgoal Interdependency Graph [46].

En cuanto a la especificación de requisitos, los autores han utilizado NFR y Kaos. La Tabla V muestra información del uso de cada modelo de objetivos de acuerdo al contexto de Big Data. Para ello se utiliza las características conocidas según la Fig. 1.

TABLA V  
MODELOS DE OBJETIVOS SEGÚN EL CONTEXTO EN BIG DATA

Título	Modelo de Objetivos	Requisitos según el contexto de Big Data
Modeling the requirements for big data application using goal oriented approach	i* y Kaos	Requisitos en función de las características propias de Big Data (volumen, velocidad, variedad, veracidad). Por lo tanto identifica aspectos técnicos del proceso de la Fig 1.
GOMA: Supporting Big Data Analytics with a Goal-Oriented Approach	Modelo Softgoal	Requisitos comerciales y de toma de decisiones. Por lo tanto, el modelo es usado para identificar actores y necesidades de análisis, que puede ser predictivo, de aprendizaje, de descubrimiento de nueva información. En los sistemas de Big Data los análisis de datos suelen ser de mayor complejidad.
The GOBIA Method: Fusing Data Warehouses and Big Data in a Goal-Oriented BI Architecture	Modelo propio (listado de objetivos)	Requisitos específicos de Inteligencia de Negocios para diseñar una arquitectura conceptual, independiente de la tecnología y centrada en la funcionalidad. El modelo permite identificar los bloques que se requieren para obtener el proceso de la Fig. 1
IRIS: A goal-oriented Big Data analytics framework on Spark for better Business decisions	Modelo de objetivos basado en softgoal de i*	Requisitos de modelo de datos en el contexto de información y análisis. En la Fig. 1 contribuye con requisitos para las etapas de almacenado, filtrado y análisis.
Big data analytics architecture design—An application in manufacturing systems	Kaos y NFR	Requisitos no funcionales para identificar los obstáculos que causan la falla del objetivo de calidad y sus tácticas de resolución correspondientes.

En el contexto de Big Data, en conjunto, los modelos de objetivos utilizados por los distintos autores permiten a los desarrolladores obtener requisitos técnicos de la arquitectura a diseñar, requisitos de análisis de datos según los objetivos definidos por el negocio, y requisitos de calidad.

No se encontraron modelos de objetivos que se utilicen para las etapas de negociación y validación. Esto sería un gran aporte debido a que los responsables de tomar decisiones tienden a identificar sus propias necesidades y no las del contexto del plan estratégico del negocio. Por otro lado, los modelos de validación permiten mejorar la calidad de la

arquitectura de un sistema.

B. P2. ¿Los modelos de objetivos utilizados por los estudios, coinciden con la clasificación en el marco de GORE propuesto por Kavakli?

Tanto NFR como Kaos han sido utilizados para tareas de especificación de requisitos, que además es lo establecido por kavakli (ver Tabla I). Sin embargo, uno de los trabajos analizados utiliza Kaos para la educación, con el fin de comparar los resultados con el uso de i\* [47].

Por otro lado, i\*, Softgoal y el modelo propio, han sido utilizados para educación de requisitos, lo que coincide con la propuesta de Kavakli.

C. P3. ¿Cuál es la contribución de los trabajos de investigación seleccionados?

Como se observa en la Fig. 3, cinco trabajos utilizan una metodología para obtener los requisitos del sistema de Big Data a través de los modelos de objetivos. Por otro lado, solo encontramos un trabajo que además utiliza una técnica.

Así mismo, solo un trabajo utiliza una herramienta de modelado para obtener los requisitos. La herramienta es basada en Spark MLlib [39] como una herramienta de asistencia de modelado (ver Fig. 4).

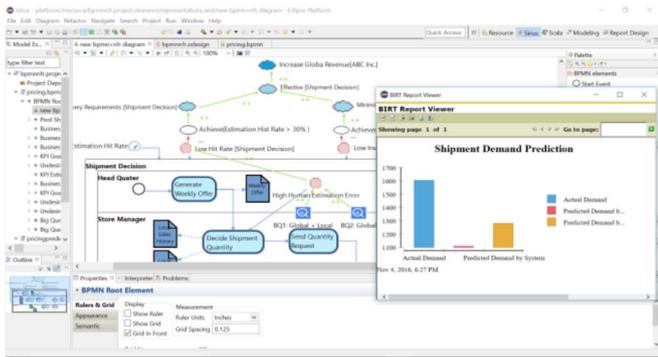


Fig. 4. Herramienta de asistencia creada por Park.

D. P4. ¿Cuál es el ámbito de aplicación de los estudios?

Tres trabajos han sido aplicados en la Academia, y sólo dos en la Industria (ver gráfico de la Fig. 5). Aún son pocos los trabajos aplicados en casos reales como para obtener conclusiones con respecto al modelado de objetivos.

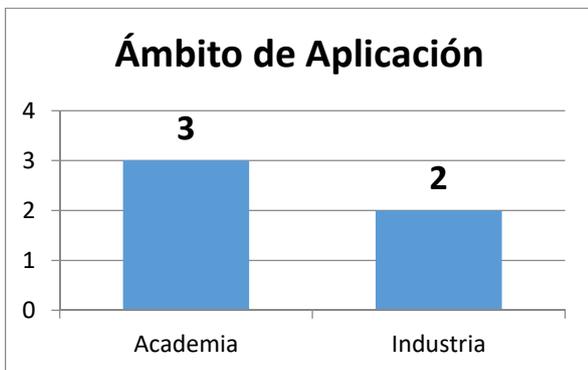


Fig. 5. Gráfico según ámbito de aplicación de los estudios seleccionados.

Por otro lado, el tipo de investigación utilizada, ya sea en la academia o industria, se presenta en la Fig. 6.



Fig. 6. Gráfico según tipo de investigación de los estudios seleccionados.

E. P5. ¿Para qué son utilizados los modelos de objetivos en sistemas de Big Data?

A continuación se describe un resumen de cada trabajo analizado:

1) Eridaputra (2014) [47]

En su trabajo propone un modelo genérico de requisitos para la aplicación Big Data a través del modelado de objetivos i\* y Kaos. El modelo de objetivos se construye a partir del análisis de los requisitos en función de las características de Big Data (volumen, velocidad, variedad, veracidad) y sus desafíos. Este modelo genérico se aplica a un estudio de caso en una agencia del gobierno de Indonesia para la planificación del desarrollo de Java Occidental (UPTB Puskalisbang Jawa Barat).

La Fig. 7 muestra el modelo de objetivos obtenidos a través de i\* para la educación de requisitos no funcionales de un sistema de Big Data, como la capacidad de la base de datos, la seguridad, la calidad y estructura de los datos, entre otros. El objetivo del modelo es mostrar los requisitos que establece Big Data de acuerdo a su definición.

Debido a que el modelo es usado para analizar los requerimientos técnicos de un sistema de Big Data, se clasifica el uso del modelo para la Implementación.

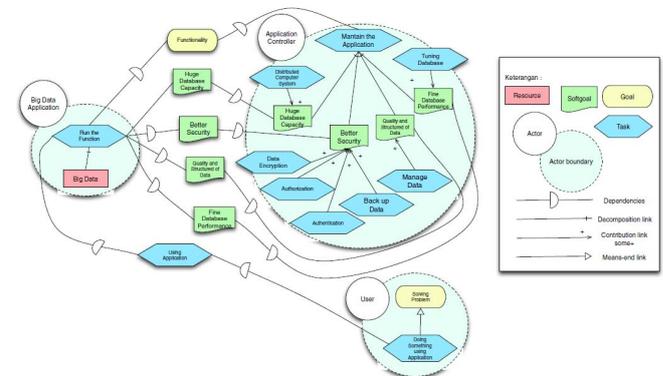


Fig. 7. Modelo SR de i\* para el sistema de Big Data según Eridaputra [47].

Los modelos i\* aprovechan las dependencias entre actores. Por otro lado, KAOS es un paradigma múltiple que permite una combinación de diferentes niveles de expresión y

razonamiento: semiformal para los modelos de objetivos, cualitativo para selecciones alternativas y formal para los elementos críticos [48].

2) *Supakkul (2016) [46]*

Según el autor, el mayor desafío al que se enfrentan las industrias no es cómo identificar los datos correctos, sino cómo utilizar los conocimientos obtenidos de Big Data para mejorar el negocio. Para abordar este desafío, los autores proponen GOMA, un enfoque de modelado orientado a objetivos para el análisis de Big Data. Basado en información de Big Data, GOMA utiliza un enfoque orientado a los objetivos para capturar los objetivos del negocio, razonar sobre situaciones comerciales y guiar los procesos de toma de decisiones. Según los autores hay dos formas de toma de decisiones: (1) conocimientos descriptivos, que son los que describen el estado actual (por ejemplo, la tasa actual de retención de clientes), y (2) conocimientos predictivos, que son los que predicen los fenómenos futuros probables por inferencia a partir de los datos (por ejemplo, los clientes que probablemente desertarán).

Con el fin ayudar en la descripción e ilustración del enfoque GOMA, los autores utilizan un caso de estudio de banca minorista.

La Fig. 8 muestra el modelo de objetivos obtenido. En el nivel más alto, se representa un objetivo de nivel corporativo, "Rentabilidad incrementada", que puede refinarse utilizando una descomposición en subobjetivos, como el "Aumento de ingresos" y "Mayores márgenes de ganancias", donde el primero debe ser operacionalizado por "Aumentar base de clientes". Luego el objetivo es descompuesto en objetivos operacionales más específicos, como "Retener clientes existentes" y "Adquirir nuevos clientes".

El modelo fue utilizado para comprender el plan estratégico del negocio, con el fin de obtener requisitos pertinentes a los responsables de tomar decisiones. En este sentido el uso del modelo sería para Extender el modelo original de  $i^*$ .

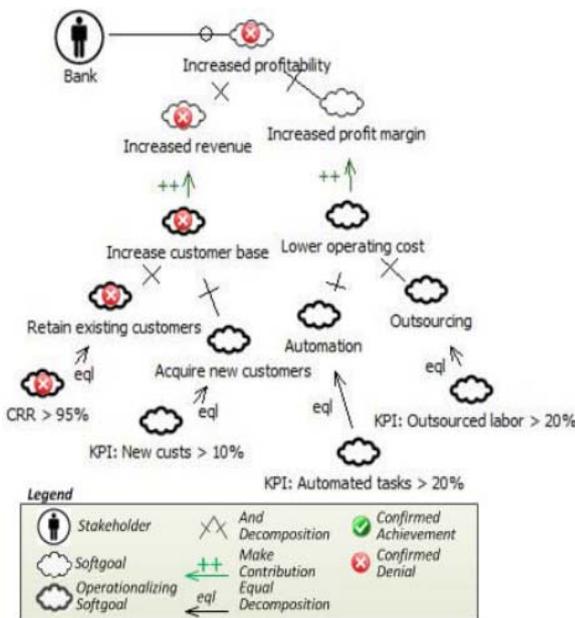


Fig. 8. Modelo de Sofgoal para el sistema de Big Data según Supakkul [46].

3) *Fekete (2016) [29]*

Los autores proponen un método llamado GOBIA que es un enfoque inicial para arquitecturas de Inteligencia de Negocios personalizadas. Los autores combinan una arquitectura de referencia de Almacenes de Datos Orientada a Objetivos existente con el proceso de desarrollo de sistemas de Big Data.

GOBIA sigue un enfoque de tres etapas, comenzando con los objetivos estratégicos del negocio, que se supone que se darán al inicio del proceso de desarrollo. Estos se utilizan para derivar requisitos específicos de Inteligencia de Negocios que serán utilizados en la segunda etapa, en la que se construye una arquitectura conceptual personalizada. Esta arquitectura conceptual, independiente de la tecnología y centrada en la funcionalidad, es la base para realizar la selección de artefactos de tecnología real y el ensamblaje de la arquitectura personalizada. En la tercera etapa, se define las tecnologías en cada bloque de la arquitectura diseñada. El modelo de objetivos que utilizan en GOBIA es propio. Consiste en un listado de objetivos estratégicos y las tareas asignadas por el negocio.

De acuerdo a la descripción del uso actual del modelado de objetivos [4], el uso en GOBIA sería para la Integración de sistemas.

4) *Park (2017) [39]*

Los autores proponen un marco de análisis orientado a objetivos para Big Data que nombran IRIS. Este marco consiste en un modelo conceptual que conecta una parte empresarial con una gran parte de datos proporcionando información de contexto alrededor de los datos, basada en método de evaluación que permite enfocar soluciones más efectivas. Los autores explican el proceso para utilizar el marco IRIS a través de un caso de estudio, y construyen una herramienta auxiliar utilizando Spark, que es una plataforma de análisis de Big Data en tiempo real. En este marco, los problemas contra los objetivos empresariales del proceso actual y las soluciones para el proceso futuro se formulan de manera explícita en el modelo conceptual y se validan con datos reales utilizando consultas o análisis de Big Data.

La Fig. 9 presenta un ejemplo de modelo de objetivos usado en IRIS. El modelo contiene un objetivo estratégico del negocio que se desglosa en subobjetivos. Luego se determinan los mecanismos de análisis posibles a utilizar y que deben ser implementados en el sistema de Big Data. El modelo ha sido utilizado para la especificación de requerimientos, por tanto su uso es para la Extensión.

5) *Fahmideh (2018)[49]*

Los autores proponen un enfoque para el análisis de objetivos y sus obstáculos para la selección de arquitecturas de Big Data que sean adecuadas para satisfacer las preferencias de objetivos de calidad y las limitaciones. El enfoque emplea modelos orientados a objetivos para identificar los obstáculos que causan la falla del objetivo de calidad y sus tácticas de resolución correspondientes. Combina la lógica difusa para explorar las incertidumbres en las arquitecturas de soluciones y para encontrar un conjunto óptimo de decisiones arquitectónicas para el proceso de habilitación de Big Data en los sistemas de fabricación. Los autores utilizan el modelo de

objetivo Kaos, como se muestra en la Fig. 10. El modelo presenta objetivos que debe cumplir el sistema de Big Data, por lo que se obtiene una especificación de requisitos funcionales, tales como “Procesado las redes sociales semanalmente en el tiempo esperado”, y “Mejora de la capacidad de la base de datos”.

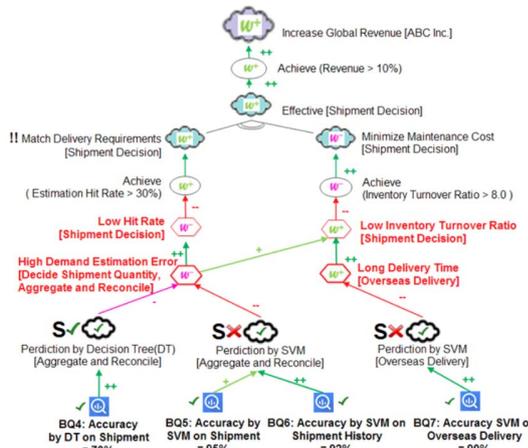


Fig. 9. Modelo de objetivos según Park [39].

El modelo de objetivos ha sido utilizado con el fin de Implementar el sistema de Big Data.

A continuación damos respuesta a las hipótesis planteadas en las preguntas de investigación.

Hipótesis 1 “GORE mejora el proceso de IR para sistemas de Big Data debido a que introduce el punto de vista del negocio en los requisitos de análisis de datos”

Hipótesis 2: “es posible adaptar los modelos de objetivos utilizados en otros contextos, en los sistemas de Big Data”

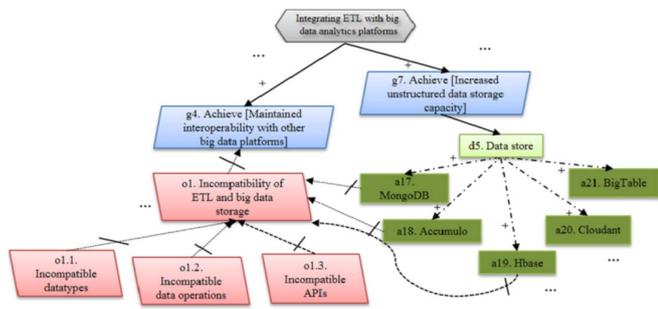


Fig. 10. Modelo de objetivos según Fahmideh [49].

La Tabla VI, presenta los resultados descritos en los trabajos seleccionados. Desde allí se pueden identificar algunos beneficios de GORE en el contexto de Big Data, como: (1) representación de los aspectos técnicos (4Vs) como metas a alcanzar, (2) identificación de KPIs de manera sistemática, (3) representación visual de los objetivos del negocio que pueden ser analizados en un sistema de Big Data, y (4) la formalización de procesos para diseñar sistemas de Big Data alineados al plan estratégico del negocio.

Desde los trabajos seleccionados, se puede diferenciar entre cuatro tipos de objetivos a nivel de GORE para el caso de sistemas de Big Data: los objetivos del negocio, los objetivos

técnicos de todo sistema de Big Data (4Vs), los objetivos a ser implementados y los objetivos que deben ser evaluados.

Otros aportes de GORE y que pueden ser beneficios para los proyectos de Big Data, se detallan en la Tabla VI. Por otro lado, estudios recientes de [4] y de [48] demuestran que GORE ha sido utilizado en distintos contextos, no solo en sistemas de apoyo a las decisiones. Los autores describen los principales beneficios para los diferentes sistemas.

No ha sido posible obtener resultados concretos del uso de GORE desde los trabajos analizados. Los autores no explican analíticamente los resultados, sino que solo son apreciaciones y opiniones de casos particulares, no generalizados a través de experimentos. En este sentido, no se puede dar una respuesta concreta para la hipótesis 1.

TABLA VI  
APORTES DE GORE PARA SISTEMAS DE BIG DATA

Beneficio de GORE según [23]	Aporte en Big Data
Los modelos de objetivos y los requisitos pueden derivarse sistemáticamente de los objetivos del negocio.	Un tema importante para los ejecutivos del negocio es que los sistemas de Big Data que se implementen deben estar alineados a los objetivos del negocio [12]. En este sentido, faltan guías, técnicas y métodos que permitan obtener requisitos de sistemas de Big Data desde los objetivos del negocio de manera sistemática [48].
Los objetivos proporcionan la justificación de los requisitos.	Los sistemas de Big Data son difíciles de implementar [49]. En este sentido, el aporte al proyecto de Big Data, es asegurar que los requisitos y necesidades obtenidos son las pertinentes al plan estratégico del negocio.
Los modelos de objetivos proporcionan una trazabilidad vertical desde las preocupaciones estratégicas de alto nivel hasta los detalles técnicos de bajo nivel.	Permite al director del proyecto de Big Data justificar los requisitos obtenidos en la primera etapa del proyecto. La trazabilidad permite identificar cada requisito con el objetivo o estrategia del negocio. Esto es importante, debido a que el diseño de la arquitectura del sistema de Big Data debe ser pertinente al procesamiento y análisis de los datos [10].
La estructura de refinamiento de objetivos proporciona una estructura comprensible para el documento de requisitos.	Los datos proporcionados por los usuarios de un sistema de Big Data son ejecutivos que toman decisiones estratégicas. En este sentido, son personas de valor estratégico para el negocio. El aporte es la comprensión de requisitos por parte de estos usuarios, lo que permite mejorar el lenguaje entre las partes interesadas y los desarrolladores. Esto es crucial, ya que en el desarrollo de sistemas de Big Data suele trabajar un equipo multidisciplinario amplio, de distintas especialidades y por lo tanto con distinto lenguaje técnico [50].

Finalmente, como respuesta a la hipótesis 2 podemos decir que los autores adaptaron los modelos de objetivos a sus propios contextos e intereses. Es el caso de los enfoques GOMA e IRIS, debido a que utilizan el modelo adaptado de i\*. Sin embargo, los autores no reportan las condiciones o

criterios que se requieren para ser adaptados, o hasta qué punto es posible modificar el modelo en el contexto de Big Data.

Por tanto, la hipótesis 2 se cumple, debido a que los autores si demuestran que es posible adaptar modelos de objetivos para el contexto de Big Data.

#### F. P6. ¿Qué resultados reportan los estudios seleccionados?

La Tabla VII presenta los resultados de cada uno de los autores.

TABLA VII  
RESULTADOS DE LOS TRABAJOS SELECCIONADOS

Autores	Resultados
<b>H. Eridaputra,</b>	El resultado de esta investigación ha demostrado que el modelo de objetivos se puede usar para generar una especificación de requisitos de software válida para ambientes de Big Data.
<b>B. Hendradjay:</b>	
<b>S. Supakkul,</b>	Los autores encontraron que el modelado de objetivos permite la identificación sistemática de KPIs importantes y útiles de una manera fácil. El documento proporciona una base inicial para vincular una serie de conceptos importantes para que se utilice de forma beneficiosa en Big Data a través de un método llamado GOMA, que proporciona un enfoque sistemático para integrar dos tipos de información, los datos y objetivos, en el proceso de toma de decisiones.
<b>L. Zhao,</b>	
<b>L. Chung</b>	
<b>D. Fekete</b>	Propuesta de una arquitectura de referencia establecida para Almacenes de Datos y Big Data que se alinean con los objetivos del negocio. Presentan un trabajo en progreso a través de un método llamado GOBIA.
<b>G. Park,</b>	
<b>S. Park,</b>	Los autores describen cuatro resultados: 1) un modelo conceptual para el análisis de requisitos que conecta el negocio con Big Data de una manera orientada a los objetivos; 2) el método de evaluación basado en la reclamación de alta prioridad llamado IRIS; 3) un proceso sobre cómo usar IRIS; y 4) una herramienta de soporte en tiempo real.
<b>L. Khan,</b>	
<b>L. Chung</b>	
<b>M. Fahmideh,</b>	El enfoque aporta dos innovaciones al estado del arte de la adopción de plataformas Big Data en sistemas de fabricación: (1) un modelo orientado a objetivos para explorar objetivos y obstáculos en la integración de sistemas de fabricación con plataformas de Big Data en el nivel de requisitos y, (2) un análisis de las decisiones arquitectónicas bajo incertidumbre.
<b>G. Beydoun</b>	

Los resultados descritos en la Tabla son solo algunos ejemplos de lo que puede aportar GORE. El aporte a los problemas que se enfrentan a la hora de implementar sistemas de Big Data puede ser aún mayor.

#### V. CONCLUSIONES

Este estudio mostró una visión general del uso de GORE en el desarrollo de sistemas de Big Data. Los estudios utilizan los modelos de objetivos i\*, Kaos, Softgoal y propios, tanto para educación como para la especificación de requisitos, no encontrándose trabajos para las etapas de negociación y validación.

Desde los trabajos analizados, podemos inferir que los modelos de objetivos han sido utilizados como modelos conceptuales para dos situaciones diferentes: (1) comprender

la naturaleza del negocio, su visión, misión, estrategias, entre otros elementos del plan estratégico con el fin de representar las necesidades de los ejecutivos en requisitos de información para el sistema de Big Data; y (2) analizar los aspectos técnicos a considerar como requisitos funcionales (volumen, velocidad, variedad, valor) y no funcionales (calidad para no fallos) del sistema de Big Data.

Sin embargo, existen muchos beneficios de GORE [6, 51], pero no todos ellos han sido utilizados en los trabajos analizados, por ejemplo: la trazabilidad, documentos de requisitos complejos, la gestión de conflictos, la validación de los requisitos encontrados, la negociación con los distintos actores y sus necesidades, entre otros. Todos estos aportes deben ser enfrentados como trabajo futuro.

Aportes concretos de GORE que pueden ser incorporados en los proyectos de Big Data son, el uso de modelos de objetivos que permitan obtener requisitos de una manera sistemática; modelos que permitan definir una trazabilidad entre los objetivos del negocio y los aspectos técnicos; y mejorar la comprensión de los requisitos técnicos por parte de los ejecutivos que utilizarán el sistema. Estos problemas se presentan en los proyectos de Big Data debido a sus características propias, como el procesamiento de volumen de información que es difícil de almacenar, y el análisis adecuado para los datos disponibles y necesidades de los ejecutivos.

#### REFERENCIAS

- [1] K. Botangen, J. Yu, S. Yongchareon, L. Yang, and Q. Bai, "Specifying and reasoning about contextual preferences in the goal-oriented requirements modelling," *Proceedings of the Australasian Computer Science Week Multiconference*, pp. 47, 2018.
- [2] P. Negri, V. Souza, A. De-Castro, R. De-Almeida, and G. Guizzardi, "Towards an Ontology of Goal-Oriented Requirements," *Coongreso Internacional en Software Engineering '17*, 2017.
- [3] E. Kavakli, "Goal Oriented Requirements Engineering: A Unifying Framework," *Requirements Engineering*, 2002.
- [4] J. Horkoff, F. Aydemir, E. Cardoso, T. Li, A. Mate, E. Paja, M. Salnitri, L. Piras, J. Mylopoulos, and P. Giorgini, "Goal-oriented requirements engineering: an extended systematic mapping study," *Requirements Engineering*, pp. 1-28, 2017.
- [5] A. Cravero and S. Sepulveda, "Using GORE in Data Warehouse: A Systematic Mapping Study," *IEEE Latin America Transactions*, vol. 13, pp. 1654--1660, 2015.
- [6] I. Reveco and A. Cravero, "GORE in the IT alignment to business strategy: A Systematic Mapping Study," *IEEE Latin America Transactions*, vol. 13, pp. 3934--3941, 2015.
- [7] J. N. Mazón, J. Pardillo, and J. Trujillo, "A Model Driven Goal Oriented Requirement Engineering Approach for Data Warehouses," *In: ER Workshops (RIGIM)*, 2007.
- [8] K. Ramasubramaniam and R. Venkatachar, "Goal-Aligned Requirements Generation," *Philips Electronics, India Limited, MFAR Manyata Tech Park, Salarpuria Infinity*, vol. 4082, 2007.
- [9] X. Jin, B. Wah, X. Cheng, and Y. Wang, "Significance and challenges of big data research," *Big Data Research*, vol. 2, pp. 59--64, 2015.
- [10] G. Park, L. Chung, L. Zhao, and S. Supakkul, "A Goal-Oriented Big Data Analytics Framework for Aligning with Business," *IEEE Third International Conference Big Data Computing Service and Applications (BigDataService)*, pp. 30--41, 2017.
- [11] M. Chen, S. Mao, and Y. Liu, "Big Data: A Survey," *Mobile Networks and Applications*, vol. 19, pp. 171--209, 2014.
- [12] R. Bean and D. Kiron, "Organizational alignment is key to big data success," *MIT Sloan Management Review*, vol. 54, pp. 1--6, 2013.
- [13] N. Madhavji, A. Miransky, and K. Kontogiannis, "Big picture of big data software engineering: with example research challenges," *Proceedings of the First International Workshop on BIG Data Software Engineering*, pp. 11--14, 2015.

- [14] A. Babar, B. Wong, and A.-Q. Gill, "An evaluation of the goal-oriented approaches for modelling strategic alignment concept.," *Conference on Research Challenges in Information Science (RCIS), 2011 Fifth International*, pp. 1-8, 2011.
- [15] S. Bleistein, K. Cox, and J. Verner, "Validating strategic alignment of organizational IT requirements using goal modeling and problem diagrams," *The Journal of Systems and Software*, vol. 79, pp. 362–378, 2006.
- [16] F. Coda, R. de Salles, F. Junqueira, D. Santos Filho, J. Silva, and P. Miyagi, "Big data systems requirements for Industry 4.0.," *13th IEEE International Conference on Industry Applications (INDUSCON)*, pp. 1230–1236, 2018.
- [17] B. Kitchenham, T. Dyba, and M. Jorgensen, "Evidence-based software engineering, in Proceeding of the 26th Int. Conf. on Software Engineering(ICSE, 2006), IEEE Computer Society.," pp. 373-378, 2006.
- [18] J.Gantz and D. Reinsel, "Extracting Value from Chaos, IDC IVIEW.," <http://www.emc.com/collateral/analystreports/idc-extracting-value-from-chaos-ar.pdf>, 2011.
- [19] Y. Demchenko, P. Membrey, P. Grosso, and C. de-Laat, "Addressing Big Data Issues in Scientific Data Infrastructure.," *First International Symposium on Big Data and Data Analytics in Collaboration (BDDAC 2013). Part of The 2013 Int. Conf. on Collaboration Technologies and Systems (CTS 2013). San Diego, California, USA*, 2013.
- [20] Y. Demchenko, C. De-Laat, and P. Membrey, "Defining architecture components of the Big Data Ecosystem," *Collaboration Technologies and Systems (CTS), 2014 International Conference*, pp. 104–112, 2014.
- [21] P. Ylijoki, "Perspectives to Definition of Big Data: A Mapping Study and Discussion," *Journal of Innovation Management JIM*, vol. 4, 2016.
- [22] P. Mikalef, I. Pappas, J. Krogstie, and M. Giannakos, "Big data analytics capabilities: a systematic literature review and research agenda," *Information Systems and e-Business Management*, vol. 16, pp. 547–578, 2018.
- [23] A. Lamsweerde, "Goal-oriented requirements engineering: A guided tour.," *Fifth IEEE International Symposium on Requirements*, 2001.
- [24] J. Horkoff, F. B. Aydemir, E. C. T. Li, A. Maté, E. Paja, M. Salnitri, J. Mylopoulos, and P. Giorgini, "Goal-Oriented Requirements Engineering: A Systematic Literature Map," *IEEE 24th International Requirements Engineering Conference*, pp. 2332-6441, 2016.
- [25] J. Horkoff, F. Aydemir, E. Cardoso, T. Li, A. Mate, E. Paja, M. Salnitri, L. Piras, J. Mylopoulos, and P. Giorgini, "Goal-oriented requirements engineering: an extended systematic mapping study," *Requirements Engineering*, vol. 24, pp. 133-160, 2019.
- [26] D. Abril and J. Pérez, "Estado actual de las tecnologías de bodega de datos y OLAP aplicadas a bases de datos espaciales," *Revista Ingeniería e Investigación*, vol. 27, 2007.
- [27] R. Winter and B. Strauch, "A method for demand driven information requirements analysis in data warehousing projects," *HICSS'03*, pp. 231-242, 2003.
- [28] R. Weir, T. Peng, and J. Kerridge, " "Best Practice for Implementing a Data Warehouse",," *A Review for Strategic Alignment, School of Computing, Napier University, 10 Colinton Road, Edinburgh EH10 5DT UK*, 2003.
- [29] D. Fekete, "The GOBIA Method: Fusing Data Warehouses and Big Data in a Goal-Oriented BI Architecture.," *GvD. GI-Workshop on Foundations of Databases (Grundlagen von Datenbanken). Nörten-Hardenberg, Germany*, pp. 50-55, 2016.
- [30] P. Eunjung, "System and method for goal-oriented big data business analytics framework," *Google Patents*, 2018.
- [31] A. Lamsweerde, R. Darimont, and P. Massonet, "Goal-Directed Elaboration of Requirements for a Meeting Scheduler: Problems and Lessons Learnt," *Requirements Engineering workshop, IEEE, York, UK.*, pp. 194-203, 1995.
- [32] E. Kavakli and P. Loucopoulos, "Experiences with goal-oriented modeling of organizational change.," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 36, pp. 221–235, 2006.
- [33] E. Kavakli and P. Loucopoulos, "Goal Driven Requirements Engineering: Evaluation of Current Methods.," *Proceedings of the 8th CAiSE/IFIP8*, vol. 1, pp. 16-17, 2003.
- [34] K. Petersen, R. Feldt, S. Mujtaba, and M. Mattsson, "Systematic mapping studies in software engineering.," *EASE'08 Proceedings of the 12th international conference on Evaluation and Assessment in Software Engineering. British Computer Society Swinton.*, pp. 68-77, 2008.
- [35] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering.," *Thechnical Report EBSE'07*, 2007.
- [36] L. Lau, F. Yang-Turner, and N. Karacapilidis, "Requirements for big data analytics supporting decision making: A sensemaking perspective.," *Mastering data-intensive collaboration and decision making*, pp. 49–70, 2014.
- [37] D. Arruda and N. H. Madhavji, "Towards a requirements engineering artefact model in the context of big data software development projects: Research in progress.," *IEEE International Conference on Big Data (Big Data)*, pp. 2314–2319, 2017.
- [38] D. Arruda, "Requirements Engineering in the Context of Big Data Applications.," *ACM SIGSOFT Software Engineering Notes*, vol. 43, pp. 1–6, 2018.
- [39] G. Park, S. Park, L. Khan, and L. Chung, "IRIS: A goal-oriented Big Data analytics framework on Spark for better Business decisions.," *IEEE International Conference on Big Data and Smart Computing (BigComp)*, pp. 76–83, 2017.
- [40] P. Clements, "Relating Business Goals to Architecturally Significant Requirements for Software Systems.," 2010.
- [41] A. Mate, J. Trujillo, E. Koci, K. Zoumpatianos, and J. Mylopoulos, "Monitoring Strategic Business Goals with Argus, Proceedings.," *IEEE International Enterprise Distributed Object Computing Workshop. EDOCW*, pp. 1-8, 2015.
- [42] NIST, "NIST Big Data interoperability Framework (NBDIF).," *NIST Big Data Use Case & Requirements Subgroup [Online]. Available*, vol. 3, 2015.
- [43] H.-M. Chen, R. Kazman, S. Haziyeve, and O. Hrytsay, "Big Data System Development: An Embedded Case Study with a Global Outsourcing Firm.," *IEEE/ACM 1st International Workshop on Big Data Software Engineering*, pp. 44–50, 2015.
- [44] Nalchigar, E. Soroosh, and Y. Eric, "Business-driven data analytics: A conceptual modeling framework.," *Data & Knowledge Engineering*, vol. 3, pp. 1–14, 2017.
- [45] C. Ardagna, P. Ceravolo, G. L. Cota, M. M. Kiani, and E. Damiani, "What are my users looking for when preparing a big data campaign.," *IEEE International Congress on Big Data (BigData Congress)*, pp. 201–208, 2017.
- [46] S. Supakkul, L. Zhao, and L. Chung, "GOMA: Supporting Big Data Analytics with a Goal-Oriented Approach.," *IEEE International Congress Big Data (BigData Congress)*, pp. 149–156, 2016.
- [47] H. Eridaputra, B. Hendradjaya, and W. D. Sunindyo, "Modeling the requirements for big data application using goal oriented approach.," *International Conference on Data and Software Engineering (ICODSE)*, pp. 1–6, 2014.
- [48] N. Kozmina, L. Niedrite, and J. Zemnickis, "Information Requirements for Big Data Projects: A Review of State-of-the-Art Approaches.," *International Baltic Conference on Databases and Information Systems*, pp. 73–89, 2018.
- [49] M. Fahmideh and G. Beydoun, "Big data analytics architecture design—an application in manufacturing systems.," *Computers & Industrial Engineering*, 2018.
- [50] A. Katal, M. Wazid, and R. Goudar, "Big data: issues, challenges, tools and good practices.," *Conference on Contemporary Computing (IC3), 2013 Sixth International*, pp. 4004–4009, 2013.
- [51] A. Lamsweerde, "Goal-Oriented Requirements Engineering: A Roundtrip from Research to Practice.," *12th IEEE International Requirements Engineering Conference (RE'04), Kyoto, Japan*, 2004.



**A. Cravero** es Licenciada en Ciencias de la Ingeniería (1996) e Ingeniera Civil Industrial m. Informática (1997), por la Universidad de La Frontera, Temuco, Chile. Obtuvo su máster en Tecnologías de la Información, por la Universidad Politécnica de Madrid, España (2006). Ha obtenido en 2010 su Doctorado en Cs. de la Computación y Sistemas Informáticos por la Atlantic International University, EE.UU.

Se desempeña como Académico en el Departamento de Ciencias de la Computación e Informática, e investigadora en

el Centro de Estudios en Ingeniería de Software, Universidad de La Frontera. Sus intereses de investigación están en el área de Modelado Bases de Datos, Ingeniería de Requisitos para Almacenes de Datos y Big Data.



**S. Sepúlveda** es Licenciado en Cs. de la Ingeniería (1998) e Ing. Civil Industrial m. Informática (1999), por la Universidad de La Frontera, Temuco, Chile. Obtuvo su máster en Dirección y Gestión de Sistemas de Información y TIC, por la Universitat Oberta de Catalunya, España (2006). Actualmente postula al grado de Doctor en Aplicaciones de la Informática por la Universidad de Alicante, España.

Se desempeña como Académico en el Dpto. de Ciencias de la Computación e Informática, e investigador en el Centro de Estudios en Ingeniería de Software, Universidad de La Frontera. Sus intereses de investigación están en el área de Ingeniería de Requerimientos, Modelado de Líneas de Productos de Software y estudios secundarios aplicados en Ingeniería de Software.